

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-162998

(43)Date of publication of application : 07.06.2002

(51)Int.Cl.

G10L 19/00

G10L 19/12

G10L 19/04

HO3M 7/30

(21)Application number : 2000-361874

(71)Applicant : FUJITSU LTD

(22)Date of filing : 28.11.2000

(72)Inventor: AMANO FUMIO

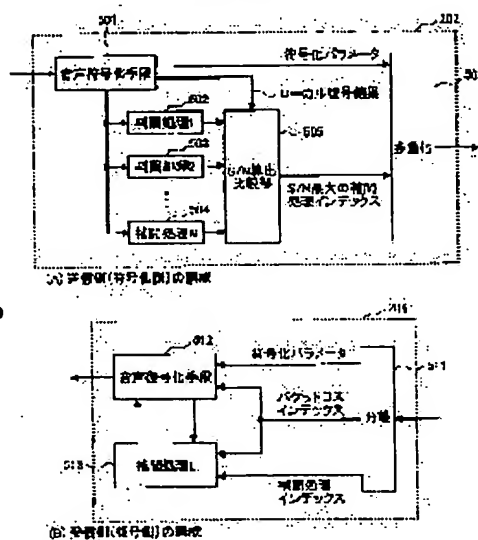
(54) VOICE ENCODING METHOD ACCOMPANIED BY PACKET REPAIR PROCESSING

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a voice encoding method accompanied by such packet repair processing that S/N and subjective quality are good and the voice in a consonant section is articulate.

SOLUTION: Multiple interpolation repair processes are prepared on a transmission side. Assuming that frames to be transmitted are lost on the transmission side, all the interpolation repair processes are tried, frame by frame. Then the waveform interpolated and repaired through the repair processes is compared with a reproduced waveform locally decoded from the packet. Then the index number of an interpolation repair processing system which can obtain an interpolated and repaired waveform closest to the locally decoded reproduced waveform is sent to a reception side together with the packet. On the reception side, multiple interpolation repair processes are prepared as well as the transmission side and if the lost of a packet is detected, a interpolation repair system is selected according to the index number of the interpolation repair system sent together with the frame and the interpolation repair process is carried out. Consequently, when no packet is lost, the interpolated and repaired waveform which is closest to the decoded reproduced waveform is obtained.

本発明の第1実施例



LEGAL STATUS

[Date of request for examination] 15.08.2002

[Date of sending the examiner's decision of rejection] 12.04.2005

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2002-162998

(P2002-162998A)

(43)公開日 平成14年6月7日(2002.6.7)

(51)Int.Cl. ⁷	識別記号	F I	テーマコード(参考)
G 1 0 L 19/00		H 0 3 M 7/30	B 5 D 0 4 5
19/12		G 1 0 L 9/18	A 5 J 0 6 4
19/04		9/14	S
H 0 3 M 7/30			J

審査請求 未請求 請求項の数 5 O L (全 11 頁)

(21)出願番号 特願2000-361874(P2000-361874)

(22)出願日 平成12年11月28日(2000.11.28)

(71)出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72)発明者 天野 文雄

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(74)代理人 100070150

弁理士 伊東 忠彦

Fターム(参考) 5D045 CC01 DA11 DA20

5J064 AA01 BA13 BB04 BC08 BC11

BC16 BC26 BD02

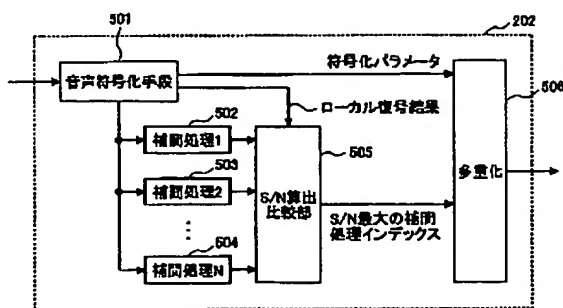
(54)【発明の名称】 パケット修復処理を伴う音声符号化方法

(57)【要約】

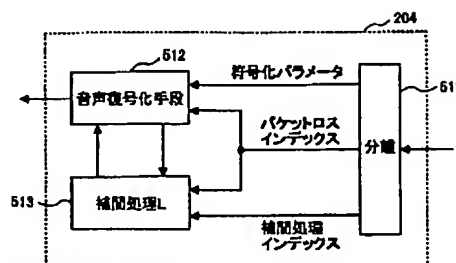
【課題】 本発明の目的は、S/Nと主観品質が良く、子音区間の音声が明瞭なパケット修復処理を伴う音声符号化方法を提供することである。

【解決手段】 送信側に複数の補間修復処理を準備する。そして、送信側で、送信する各フレーム毎に、そのフレームが消失したと仮定して、この全ての補間修復処理を試みる。そして、修復処理を行って補間修復した波形と、そのパケットからローカルに復号した再生波形との比較を行う。この結果、ローカルに復号した再生波形に最も近い補間修復した波形が得られる補間修復処理方式のインデックス番号を受信側にパケットと共に送信する。受信側では、送信側と同様に、複数の補間修復処理を準備し、パケットの消失を検出した場合には、そのフレームと共に伝送される補間修復方式のインデックス番号に従って、補間修復処理方式を選択し、補間修復処理を行う。これにより、パケットが消失しなかった場合に復号した再生波形に最も近い補間修復した波形が得られる。

本発明の第1実施例



(A) 送信側(符号化側)の構成



(B) 受信側(復号側)の構成

【特許請求の範囲】

【請求項 1】 音声信号を短時間に区切って、音声パラメータを抽出して音声フレームとする手段と、

現在の音声フレームを基に、第 1 の音声に再生する手段と、

前記現在の音声フレーム以外の音声フレームを用いて、複数の補間処理して得られる複数の音声フレームを生成する手段と、

該音声フレームを基に、複数の第 2 の音声に再生する手段と、

該第 2 の音声のうち、該第 1 の音声に近い該第 2 の音声に該当する補間処理を示す識別情報を出力する決定手段と、

前記現在の音声フレームに、該識別情報を多重化して送信する多重化手段とを有することを特徴とする音声符号化器。

【請求項 2】 複数の音声データを有する第 1 のフレームを、符号化パラメータに符号化する符号化ステップと、

前記第 1 のフレーム内に子音を含むか否かを検出するステップと、

前記検出するステップにより、前記第 1 のフレームが子音を含む場合には、前記第 1 のフレームに同一のシーケンス番号を付加した同一のフレームを、複数回送信するステップとを有する音声符号化方法。

【請求項 3】 複数の音声データを有する第 1 のフレームを、符号化パラメータに符号化する符号化ステップと、

前記第 1 のフレーム内に子音を含むか否かを検出するステップと、

前記検出するステップにより、前記第 1 のフレームが子音を含む場合には、前記第 1 のフレームに高い優先度を示す情報を付加して送信するステップとを有する音声符号化方法。

【請求項 4】 複数の音声データを有する第 1 のフレームを、符号化パラメータに符号化する符号化ステップと、

前記第 1 のフレームの符号化されたパラメータを局部的に第 2 のフレームに復号するステップと、

前記第 1 のフレーム以外のフレームを用いて、前記第 1 のフレームの近似フレームを生成する複数の補間修復処理を行うステップと、

前記複数の補間修復処理を行うステップの各々により生成された前記第 1 のフレームの近似フレームと、前記第 2 のフレームとを比較し、各々の前記第 1 のフレームの近似フレームに対して、前記第 2 のフレームを信号として信号対雑音比を計算し、且つ、前記信号対雑音比が最も高くなる前記複数の補間修復処理のうちの 1 つを示すインデックス番号を決定するステップと、

前記第 1 のフレーム内に子音を含むか否かを検出するス

テップと、

前記検出するステップにより、前記第 1 のフレームが子音を含む場合には、前記決定するステップにより決定された前記信号対雑音比が最も高くなる前記複数の補間修復処理のうちの 1 つを示すインデックス番号を、前記符号化パラメータと共に多重して更に、同一のシーケンス番号を付加して複数回送信するステップとを有する音声符号化方法。

【請求項 5】 複数の音声データを有する第 1 のフレームを、符号化パラメータに符号化する符号化ステップと、

前記第 1 のフレームの符号化されたパラメータを局部的に第 2 のフレームに復号するステップと、

前記第 1 のフレーム以外のフレームを用いて、前記第 1 のフレームの近似フレームを生成する複数の補間修復処理を行うステップと、

前記複数の補間修復処理を行うステップの各々により生成された前記第 1 のフレームの近似フレームと、前記第 2 のフレームとを比較し、各々の前記第 1 のフレームの近似フレームに対して、前記第 2 のフレームを信号として信号対雑音比を計算し、且つ、前記信号対雑音比が最も高くなる前記複数の補間修復処理のうちの 1 つを示すインデックス番号を決定するステップと、

前記第 1 のフレーム内に子音を含むか否かを検出するステップと、

前記検出するステップにより、前記第 1 のフレームが子音を含む場合には、前記決定するステップにより決定された前記信号対雑音比が最も高くなる前記複数の補間修復処理のうちの 1 つを示すインデックス番号を、前記符号化パラメータと共に多重して、更に高い優先度を示す情報を付加して送信するステップとを有する音声符号化方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、IP（インターネットプロトコル）ネットワークを利用して音声の伝送を行うための音声符号化方法に関し、特に伝送中にパケットが消失した場合に、受信側での再生音声品質の劣化を軽減することが可能な音声符号化方法に関する。

【0002】

【従来の技術】IPネットワークを利用して音声の伝送を行う技術としてVOIP（Voice Over IP）が知られている。図1は、VOIP伝送系の基本構成を示す。VOIP伝送系は、主に、電話機等のユーザ端末101、107、アクセス系/既存網102、106、VOIPGW（VOIPゲートウェイ）103、105及び、インターネット104より構成される。VOIPGW103、105は、アクセス系/既存網102、106とインターネット104の間に配置される。

図2は、VOIPGW音声処理部の基本構成を示す。V

3

VOIPGW音声処理部は、主に、アクセス系/既存網インターフェース201、音声符号化部202、パケット組立部203、音声復号化部204、パケット分解部205より構成される。VOIPは、アクセス系/既存網102、106を介してVOIPGW103、105に

入力される音声を、音声符号化部202で低ビットレートで符号化して伝送し、データパケット伝送と混在させることにより音声通話の低コスト化を図るものである。
【0003】しかし、図1の基本構成では、例えば、以下に示すような問題がある。第1は、パケットが、IPネットワーク上で複数のルータを経由されて伝送されるために、遅延が大きくなることである。第2は、パケットが各種バッファを経由して伝送されることにより、パケットが受信側に到着する時間にゆらぎ（ジッタ）を生じることである。第3は、各種バッファでデータのオーバーフローが生じたり、又は、伝送中にエラーが生じることによりパケットの消失が発生して、受信側での再生音声品質が劣化することである。

【0004】消失したパケットを補償するための送信側の従来技術としては、例えば、次のような技術がある。第1は、パケット消失情報を受信側から送信側へ送り返し、対象フレームを再送する方式である。第2は、インターリーブ処理を使用して誤りをランダムとして、パケット消失の影響を軽減する方式である。第3は、FEC（フォワードエラーコレクション）符号化を行う方式である。

【0005】また、受信側の従来技術としては、例えば、次のような技術がある。第1は、消失フレームに対して波形の挿入を行う方式である。第2は、消失フレーム前後のフレーム又は、消失フレーム前のフレームの波形から波形の内挿補間を行う方式である。第3は、消失フレーム前後のフレームの音声符号化パラメータを内挿補間し、補間したパラメータから音声を再生する方式である。これらの技術は、1998年9月/10月の、'A Survey of Packet Loss Recovery Techniques for Streaming Audio' IEEE Network Magazine、40から18頁、及び、2000年4月の'Internet Telephony: Services Technical Challenges, and Products'、IEEE Communication Magazine、96から103頁に記載されている。

【0006】上記の送信側の第1及び第2の従来技術は、遅延時間等が大きくても許容されるいわゆる配信サービスにおいて主に使用される。図3は、上記の送信側の第3の従来技術のメディア特有の補間処理の例を示す。

【0007】図3において、参照番号301から304はオリジナル音声ストリームの各フレームを示す。本例

4

の場合は、4フレームが示されている。本例においては、例えば、フレーム303を符号化する場合には、通常使用する符号化パラメータ313-3と、通常使用するよりは低ビットレートの音声符号化器の符号化パラメータ314-3の2種類のパラメータへの符号化を行う。そして、通常使用する符号化パラメータ313-3はフレーム313で、また、通常使用するよりは低ビットレートの音声符号化器の符号化パラメータ314-3はフレーム314で、それぞれFECを付加してパケット化して伝送する。伝送中に、例えば、パケット313が消失した場合には、再生側において、通常使用する符号化パラメータ313-3の代わりに、通常使用するよりは低ビットレートの音声符号化器の符号化パラメータ314-3を使用して、パケット313で伝送されるべきであった音声フレーム303に対応する音声波形を再生する。この方式の処理遅延時間は1フレーム期間であり、そして、ある程度以上の音声品質を得るためには、低ビットレートの符号化器として2から4k bps程度で符号化できる符号化器が必要である。従って、低ビットレートの音声符号化器の符号化パラメータ314-3を付加するには、フレーム長が20msecの場合、40から80ビットの冗長なデータ（オーバーヘッド）が必要である。

【0008】これに対して、上記の消失パケットを受信側で補間する従来技術を使用することにより、オーバーヘッドを設けずに補間処理を行うことができる。図4は、受信側における従来の補間処理方式の基本構成を示す図である。図4は、図2の音声復号化部204を示す。図4では、音声復号化部204は、主に、パケット分離部401、音声復号化部402、補間処理部403よりなる。パケット分離部401から出力される符号化パラメータは、音声復号化部402に与えられ、音声波形が再生されそして、出力される。一方、受信パケットに消失がある場合には、消失したパケットを示すパケットロスインデックスにより、パケットに消失があることが補間処理部403に通知され、補間処理部403において消失フレームの補間処理が行われる。補間処理は、例えば、次のように行われる。

【0009】第1は、パケットの消失（パケットロス）の生じた前のフレームの再生した波形に、窓関数を乗じて、その波形を、パケットロスの生じたフレームの再生波形として使用する。或は、第2に、符号化パラメータを、パケットロスの生じたフレームの前後又は、前のフレームから補間し、補完されたパラメータを使用して、パケットロスの生じたフレームの音声を再生することも可能である。この場合、例えば、LPC（線形予測符号化）パラメータについては、パケットロスの生じたフレームの前後のフレームのパラメータから、パラメータの線形補間を行う。それ以外のパラメータについては、パケットロスの生じたフレームの前のフレームのパラメータ

5

タと同じ値のパラメータを使用する。

【0010】

【発明が解決しようとする課題】受信側において消失パケットを補間し修復する処理においては、パラメータの補間を行う方式が、再生品質の維持を図ることに関して優れていることが知られている。しかし、この方式には、次のような問題点がある。

【0011】第1は、補間修復処理方式には複数の方式があるが、従来の方式では、特定の1つの方式のみに従って処理を行う構成となっている。これにより、S/N（信号対雑音比）又は、主観品質の観点からは、必ずしも最適な方式で、消失パケットの補間修復処理が行われてはいない。

【0012】第2は、消失したフレームに子音区間が含まれている場合には、補間修復処理を行っても、音声の明瞭性が失われてしまうことである。

【0013】

【課題を解決するための手段】本発明は、上記課題を解決した、S/Nと主観品質が良く、子音区間の音声が見えやすいパケット修復処理を伴う音声符号化方法を提供することを目的とする。

【0014】上記第1の課題に対しては、送信側に複数の補間修復処理を準備する。そして、送信側で、送信する各フレーム毎に、そのフレームが消失したと仮定して、この全ての補間修復処理を試みる。そして、修復処理を行って補間修復した波形と、そのパケットからローカルに復号した再生波形との比較を行う。この結果、ローカルに復号した再生波形に最も近い補間修復した波形が得られる補間修復処理方式のインデックス番号を受信側にパケットと共に送信する。受信側では、送信側と同様に、複数の補間修復処理を準備し、パケットの消失を検出した場合には、そのフレームと共に伝送される補間修復方式のインデックス番号に従って、補間修復処理方式を選択し、補間修復処理を行う。これにより、パケットが消失しなかった場合に復号した再生波形に最も近い補間修復した波形が得られる。

【0015】一方、上記第2の課題に対しては、送信側で、各フレーム毎に、そのフレームが子音区間を含むか否かの検出処理を行う。そして、子音を含む場合には、そのフレームの優先度を高くして送信する。優先度を高くするには、例えば、子音を含むフレームを複数回送信するか又は、フレームの優先度の設定が可能な場合には子音を含むフレームの優先度を高く設定する等の処理を行う。

【0016】

【発明の実施の形態】本発明は、図1におけるVOIP GW103、105に適用される。図5は、本発明の第1実施例を示す図であり、本実施例は、上記第1の課題を解決するための基本構成を示す。図5(A)は、図2の送信側としての構成である音声符号化部202の構成

6

の例を示す。また、図5(B)は、図2の受信側としての構成である音声復号化部204の構成の例を示す。音声符号化部202は、主に、音声符号化手段501、補間処理部502、補間処理部503、補間処理部504のような複数の補間処理部、S/N算出比較部505、多重化部506を有する。音声符号化部501は、符号化を行った結果の符号化パラメータから、符号化器内で局部的に復号を行うローカル復号部も有する。このローカル復号部は、符号化器の一部として構成されているものを使用することも可能である。また、音声復号化部204は、分離部511、音声復号化手段512、補間処理部513を有する。送信側では、補間処理部502、補間処理部503、補間処理部504では、各フレーム毎に、そのフレームが消失したと仮定して、各補間修復処理を試みる。そして、補間処理部502、補間処理部503、補間処理部504が、それぞれ、修復処理を行って補間修復した波形と、音声符号化手段501によりそのパケットからローカルに復号した再生波形のS/Nを、S/N算出比較部505により比較する。この結果、最も高いS/Nが得られる補間処理部に対応する補間修復処理方式のインデックス番号と符号化パラメータが多重化部506に送られ、多重されて送信される。一方、受信側では、パケット消失が無い場合には、分離部511から出力される符号化パラメータを用いて、音声復号化手段512により音声復号処理を行う。分離部511でパケットの消失が検出された場合には、送信された補間修復処理方式のインデックス番号を用いて、補間修復処理を行う。

【0017】図6は、図5に示す本発明の第1実施例の処理の流れを示す。図6(A)は、入力音声信号の各フレーム601、602、603を示し、(B)は各処理の期間611から616を示し、(C)は出力されたパケット621、622、623及び、パケット622の構成例を示す。また図6(D)はパケットの消失がない場合の受信側で受信されたパケット631、632、633と、その復号された音声出力641、642、643を示す。また、図6(E)はパケットの消失がある場合の受信側で受信されたパケット631、632、633と、その復号された音声出力641、644、643を示す。

【0018】送信側においては、音声入力フレーム601、602、603は、処理期間611、612、613で、音声符号化処理が行われる。一方、処理期間614、615、616では、上述した、補間処理部502、503、504等により、各フレーム毎に、そのフレームが消失したと仮定して、各補間修復処理を行う。例えば、処理期間616では、フレーム602に対して、フレーム601と603の符号化パラメータを用いて、各補間修復処理を行い、最も高いS/Nが得られる補間処理部に対応する補間修復処理方式のインデックス

7

番号を算出する。そして、この算出したインデックス番号を、符号化パラメータと共にバケット化する。バケットは、例えば、ヘッダ部625、制御ビット626、算出した最適な補間処理の方式のインデックス番号627、符号化パラメータ628より構成される。図7は、他のバケットの構成例を示したものである。例えば、バケットは、IPヘッダ701、UDPヘッダ702、RTPヘッダ703及び、音声符号化データ704より構成される。上述の算出したインデックス番号は、例えば、IPヘッダ701中のTOS（サービスタイプ）フィールド705の、ビット6と7のような未使用領域に配置されても良い。このように、バケットの符号化データ704の領域外に、インデックス番号を配置することにより、音声品質を損なうことなく、インデックス番号を送信することが可能である。また、同様に、例えば、RTPヘッダ703に未使用領域がある場合には、そこへ配置することも可能である。さらに、符号化データ704の領域中には、エラーに対する感度の低い領域もあるので、エラーに対する感度の最も低い領域に、算出したインデックス番号を配置する構成とすることにより、音声品質に与える影響を最小にしてインデックス番号を符号化データ704の領域中に配置して送信することも可能である。

【0019】また、この符号化データ704の領域中のエラーに対する感度の最も低い領域に算出したインデックス番号を配置することにより、インデックス番号を符号化データ704の領域中に配置して送信する方法においては、インデックス情報を数フレームに対して1回伝送することにより、更に音声品質の劣化を抑えることができる。この場合には、数フレームに1回の割合で上述の処理を行い、或は、例えば、隣接するフレーム間の音声符号化パラメータの変化が大きい場合にのみ前述の処理を行い、インデックス番号を算出して送信することができる。

【0020】一方、受信側においては、図6(D)に示すように、バケットの消失がない場合には、受信されたバケット631、632、633から、そのフレームの符号化パラメータを用いて音声出力641、642、643が復号される。一方、図6(E)に示すように、例えば、バケット632が消失している場合には、フレーム631と633の符号化パラメータと同時に送信されたインデックス番号を用いて補間修復処理を行い、音声フレーム644を再生する。

【0021】次に、本発明の第2実施例について説明する。図8(A)は、音声符号化方法にCELP方式を使用した場合の実施例である。図8(A)の音声符号化部202は、CELP符号器801、フレームバッファ802、803、804、補間処理部805、806、807、808、ローカル復号部809、810、811、812、S/N算出比較部813及び、多重化部8

8

14を有する。また、図9は、CELP符号化器801の構成を示す。CELP符号器801は、主にLPC分析部901、LPC量子化部902、合成フィルタ部903、減算部904、聴覚重み付けフィルタ部905、歪最小化部906、適応符号帳907、固定符号帳908、ゲイン調整部909、910及び、加算部811を有する。

【0022】CELP方式は、AbS(Analysis by Synthesis、分析による合成)を行うことにより、最適な符号帳を選択することにより音声の圧縮を行う方式である。CELP符号器801では、例えば20msecのフレームごとにLPCパラメータをLPC分析部901で算出し、また例えば、5msecのサブフレームごとに最適な音声品質が得られる適応符号帳のインデックスとゲイン及び、固定符号帳のインデックスとゲインを算出し出力する。図9(B)は、フレームとサブフレームの関係を示す。図8(A)では、CELP符号器801で算出された上記の各パラメータは、2フレーム前の値までフレームバッファ802に蓄積される。同様に、ローカル復号部の内部状態及び合成フィルタ903の出力は、1フレーム前の値がフレームバッファ803に蓄積される。そして、各フレーム毎に、1フレーム前のフレームが伝送により消失したと仮定して、各補間処理部805から808で補間修復処理を行う。

【0023】図8(A)の補間処理805においては、LPCパラメータについて、2フレーム前の値と現フレームの値を用いて線形補間処理を行う。適応符号帳インデックスとゲイン及び、固定符号帳インデックスとゲインについては、4つのサブフレームの全てについて、2フレーム前の第4番目のサブフレームの値をそのまま用いる。

【0024】図8(A)の補間処理806においては、LPCパラメータについて、補間処理805と同様に、線形補間処理を行う。適応符号帳インデックスとゲイン及び、固定符号帳インデックスとゲインについては、第1番目のサブフレームには、2フレーム前の第3番目のサブフレームの値を、第2番目のサブフレームには、2フレーム前の第4番目のサブフレームの値を、第3番目のサブフレームには、現在のフレームの第1番目のサブフレームの値を、そして、第4番目のサブフレームには、現在のフレームの第2番目のサブフレームの値をそれぞれ使用する。

【0025】図8(A)の補間処理807においては、LPCパラメータの補間については、2フレーム前の値と現在のフレームの値から、2次関数補間処理を行う。その他のパラメータに関しては、補間処理805と同様な処理を行う。

【0026】図8(A)の補間処理808においては、LPCパラメータの補間については、2フレーム前の値

10

20

30

40

50

と現在のフレームの値から、2次関数補間処理を行う。その他のパラメータに関しては、補間処理806と同様な処理を行う。以上のような4つの補間処理により得られたパラメータを用いて、ローカル復号部809、810、811、812で、それぞれローカル復号を行う。そして、1フレーム前の符号化パラメータを用いたローカル復号の出力と、ローカル復号部809、810、811、812の出力とをS/N算出比較部813において比較し、S/N値を算出する。そして、S/N値が最も大きくなる補間方式を選択し、そのインデックス情報をCELP符号化パラメータと共に多重化部814により多重して、パケット組立部203へ送る。

【0027】例えば、補間処理部805、806、807、808の各処理に、インデックス番号00、01、10、11を対応させる。そして、補間処理部807の出力から最も高いS/N値が得られる場合には、10をインデックスとして多重する。

【0028】以上説明した処理は、例えば、DSP（デジタル信号処理プロセッサ）のファームウェア処理により実現することができる。

【0029】図8（B）は、復号器側の構成を示す。音声復号化部204は、パケット分離部821、フレームバッファ822、補間処理部823、選択器824及び、CELP復号器825を有する。受信された符号化パラメータは、パケット分離部821で分離され1フレーム分のフレームバッファ822へ蓄積される。同時に送信されたパケットロスインデックスによりフレームの消失が通知された場合には、補間処理部823によりインデックスの示す最適な補間処理を選択し補間修復処理を行う。

【0030】次に本発明の第3実施例について説明する。図10は、本発明の第3実施例を示す図であり、図2の音声符号化部202とパケット組立部203の構成の実施例を示す。音声符号化部202は、音声符号化手段1001と、母音/子音検出部1002を有する。入力された音声は、フレーム毎に、音声符号化手段1001で符号化されると共に、母音/子音検出部1002によりそのフレームに子音期間が含まれるか否かが検出される。そして、子音期間の検出結果は、符号化パラメータと共にパケット組立部203へ送られる。パケット組立部203では、そのフレームに子音期間を含む場合には、パケット送出バッファの充填度を観測しながら、次のフレームの処理が行われる前に、複数回にわたって同一のシーケンス番号を付加して同一フレームを送出する。

【0031】図11は、本発明の第3実施例の処理の流れを示す図である。図11（A）は、入力音声信号の各フレーム1101、1102、1103を示し、（B）は各処理の期間1111から1116を示し、（C）は出力されたパケット1121、1122、1123、1

124、1125を示す。また図11（D）は子音期間を含むパケットが消失した場合の受信側で受信されたパケット1121から1125と、その復号された音声出力1131、1132、1133を示す。

【0032】送信側では、図11（A）において入力された各音声フレームに対して、図11（B）において、処理期間1111、1112、1113で音声符号化手段1001で符号化されると共に、処理期間1114、1115、1116で母音/子音検出部1002によりそのフレームに子音期間が含まれるか否かが検出される。例えば、フレーム1102に子音期間を含むことが検出された場合には、パケット組立部203では、パケット送出バッファの充填度を観測しながら、次のフレーム1103の処理が行われる前に、複数回にわたって同一のシーケンス番号を付加して同一フレーム1122、1123、1124を送出する。

【0033】一方、受信側においては、パケット1121を受信した後、次に予想される時間に次のパケット1122が受信できなかった場合には、パケットの消失の可能性を考慮して、複数回にわたって同一のシーケンス番号を付加して同一フレームが送出された時間の間、パケットの受信を待つ。そして、その間に、例えば、同一のシーケンス番号が付されたパケット1123が受信された場合には、そのパケットによりフレーム1132の復号を行う。

【0034】次に本発明の第4実施例について説明する。図12は、本発明の第4実施例を示す図である。図12（A）は送信側の構成を示し、主に音声符号化部202とパケット組立部203で構成される。音声符号化部202は、CELP符号化部1201、零交差数検出部1202、logレベル検出部1203、1次自己相関値検出部1204、子音期間検出部1205を有する。図12（B）は、零交差数Z、logレベルL、及び、1次自己相関値Rの分布例を示す。本実施例においては、対象となるフレームに対して、サブフレーム毎に、子音期間検出部1205により子音期間検出を行う。子音期間検出は、サブフレーム毎に、零交差数Z、logレベルL、及び、1次自己相関値Rを算出する。そして、算出したこれらの値と、零交差数の所定のしきい値Thz、logレベルの所定のしきい値Thl、及び、1次自己相関値の所定のしきい値Thrと比較する。Z>Thz、L<Thl、及び、R>Thrが同時に成立した場合にはそのサブフレームを子音期間と判定する。そして、対象フレームの中に、1つでも子音期間のサブフレームがあれば、そのフレームを子音期間と判定する。なお、母音、子音、無音の各区間を識別する方法としては、例えば、1976年7月の'A Pattern Recognition Approach to Voiced-Unvoiced-Silence Classification with Appl

ications of Speech Recognition!」、IEEE Trans. on ASSP, ASSP-24, No. 3, 201から212頁に記載されている。本実施例では、上記論文の図2, 3, 4の性質を利用する方式を採用している。

【0035】図12(C)は、受信側の構成を示す。受信側は、フレームバッファ1211、パケット分解部1212、CELP復号化部1213を有する。フレームバッファ1211により、パケット消失の可能性を考慮して、送信側で複数回にわたって同一のシーケンス番号を付加して同一フレーム送出されたフレームまでの時間の間、パケットの受信を待ち、例えば、同一のシーケンス番号が付されたパケットが受信された場合には、そのパケットによりフレームの復号を行う。図12の全体の処理は、例えば、DSP（ディジタル信号処理プロセッサ）のファームウェア処理により実現することができる。

（付記）

（付記1） 音声信号を短時間に区切って、音声パラメータを抽出して音声フレームとする手段と、現在の音声フレームを基に、第1の音声に再生する手段と、前記現在の音声フレーム以外の音声フレームを用いて、複数の補間処理して得られる複数の音声フレームを生成する手段と、該音声フレームを基に、複数の第2の音声に再生する手段と、該第2の音声のうち、該第1の音声に近い該第2の音声に該当する補間処理を示す識別情報を出力する決定手段と、前記現在の音声フレームに、該識別情報を多重化して送信する多重化手段とを有することを特徴とする音声符号化器。

【0036】（付記2） 前記第1のフレーム以外のフレームは、前記第1のフレームよりも前のフレームである付記1記載の方法。

【0037】（付記3） 前記第1のフレーム以外のフレームは、前記第1のフレームよりも前のフレームと後のフレームである付記1記載の方法。

【0038】（付記4） 前記送信するステップは、前記信号対雑音比が最も高くなる前記複数の補間修復処理のうちの1つを示すインデックス番号を、パケット内の符号化パラメータを配置する領域以外の領域に配置することにより、前記インデックス番号を送信する付記1記載の音声符号化方法。

【0039】（付記5） 前記送信するステップは、前記信号対雑音比が最も高くなる前記複数の補間修復処理のうちの1つを示すインデックス番号を、パケット内の符号化パラメータを配置する領域内のエラーに対する感度の最も低い領域に配置することにより、前記インデックス番号を送信する付記1記載の音声符号化方法。

【0040】（付記6） 複数の音声データを有する第1のフレームを、符号化パラメータに符号化する符号化ステップと、前記第1のフレーム内に子音を含むか否か

を検出するステップと、前記検出するステップにより、前記第1のフレームが子音を含む場合には、前記第1のフレームに同一のシーケンス番号を付加した同一のフレームを、複数回送信するステップとを有する音声符号化方法。

【0041】（付記7） 複数の音声データを有する第1のフレームを、符号化パラメータに符号化する符号化ステップと、前記第1のフレーム内に子音を含むか否かを検出するステップと、前記検出するステップにより、前記第1のフレームが子音を含む場合には、前記第1のフレームに高い優先度を示す情報を付加して送信するステップとを有する音声符号化方法。

【0042】（付記8） 複数の音声データを有する第1のフレームを、符号化パラメータに符号化する符号化ステップと、前記第1のフレームの符号化されたパラメータを局部的に第2のフレームに復号するステップと、前記第1のフレーム以外のフレームを用いて、前記第1のフレームの近似フレームを生成する複数の補間修復処理を行うステップと、前記複数の補間修復処理を行うステップの各々により生成された前記第1のフレームの近似フレームと、前記第2のフレームとを比較し、各々の前記第1のフレームの近似フレームに対して、前記第2のフレームを信号として信号対雑音比を計算し、且つ、前記信号対雑音比が最も高くなる前記複数の補間修復処理のうちの1つを示すインデックス番号を決定するステップと、前記第1のフレーム内に子音を含むか否かを検出するステップと、前記検出するステップにより、前記第1のフレームが子音を含む場合には、前記決定するステップにより決定された前記信号対雑音比が最も高くなる前記複数の補間修復処理のうちの1つを示すインデックス番号を、前記符号化パラメータと共に多重して更に、同一のシーケンス番号を付加して複数回送信するステップとを有する音声符号化方法。

【0043】（付記9） 前記第1のフレーム以外のフレームは、前記第1のフレームよりも前のフレームと後のフレームである付記8記載の方法。

【0044】（付記10） 複数の音声データを有する第1のフレームを、符号化パラメータに符号化する符号化ステップと、前記第1のフレームの符号化されたパラメータを局部的に第2のフレームに復号するステップと、前記第1のフレーム以外のフレームを用いて、前記第1のフレームの近似フレームを生成する複数の補間修復処理を行うステップと、前記複数の補間修復処理を行うステップの各々により生成された前記第1のフレームの近似フレームと、前記第2のフレームとを比較し、各々の前記第1のフレームの近似フレームに対して、前記第2のフレームを信号として信号対雑音比を計算し、且つ、前記信号対雑音比が最も高くなる前記複数の補間修復処理のうちの1つを示すインデックス番号を決定するステップと、前記第1のフレーム内に子音を含むか否か

13

を検出するステップと、前記検出するステップにより、前記第1のフレームが子音を含む場合には、前記決定するステップにより決定された前記信号対雑音比が最も高くなる前記複数の補間修復処理のうちの1つを示すインデックス番号を、前記符号化パラメータと共に多重して、更に高い優先度を示す情報を付加して送信するステップとを有する音声符号化方法。

【0045】

【発明の効果】以上、本発明により、S/Nと主観品質が良く、子音区間の音声が見え難いパケット修復処理を伴う音声符号化方法を提供することができる。

【図面の簡単な説明】

【図1】VOIP伝送系の基本構成を示す図である。

【図2】VOIPGW音声処理部の基本構成を示す図である。

【図3】送信側での従来のメディア特有の補間処理の例を示す図である。

【図4】従来の補間処理方式の基本構成を示す図である。

【図5】本発明の第1実施例を示す図である。

【図6】本発明の第1実施例の処理の流れを示す図である。

【図7】パケットの構成例を示す図である。

【図8】本発明の第2実施例を示す図である。

【図9】CELP符号化方式を示す図である。

【図10】本発明の第3実施例を示す図である。

【図11】本発明の第3実施例の処理の流れを示す図である。

【図12】本発明の第4実施例を示す図である。

【符号の説明】

101、107 電話機等のユーザ端末

14

102、106 アクセス系/既存網

103、105 VOIPGW

104 インターネット

201 アクセス系/既存網インターフェース

202 音声符号化部

203 パケット組立部

204 音声復号化部

205 パケット分解部

401 パケット分離部

10 402 音声復号化部

403 補間処理部

501 音声符号化手段

502、503、504 補間処理部

505 S/N算出比較部

506 多重化部

511 分離部

512 音声復号化手段

513 補間処理部

801 CELP符号器

20 802、803、804 フレームバッファ

805、806、607、808 補間処理部

809、810、811、812 ローカル復号部

813 S/N算出比較部

814 多重化部

821 パケット分離部

822 フレームバッファ

823 補間処理部

824 選択器

825 CELP復号器

30 1001 音声符号化手段

1002 母音/子音検出部

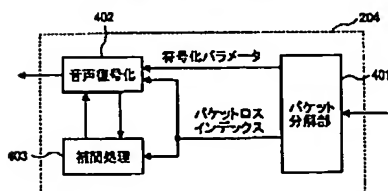
【図1】

VOIP伝送系の基本構成



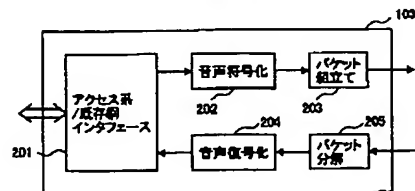
【図4】

従来方式の基本構成

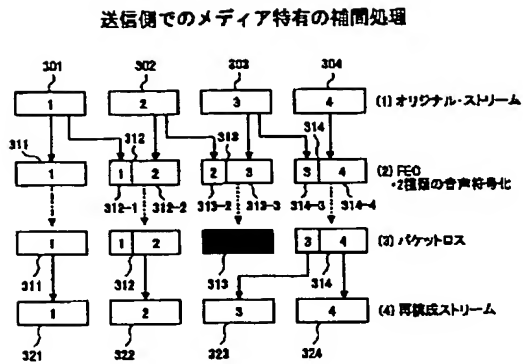


【図2】

VOIPGW音声処理部の基本構成

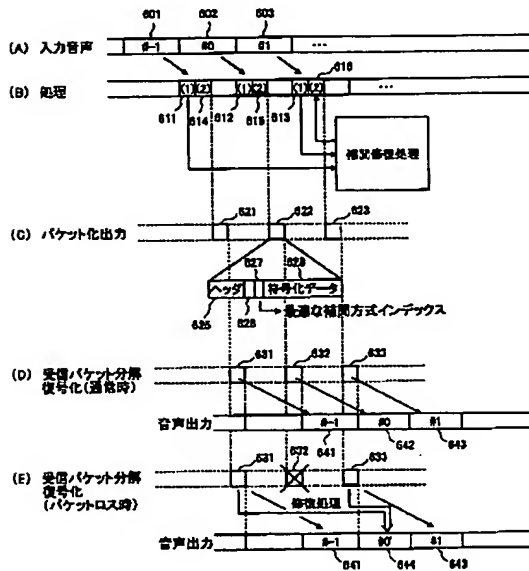


【図 3】



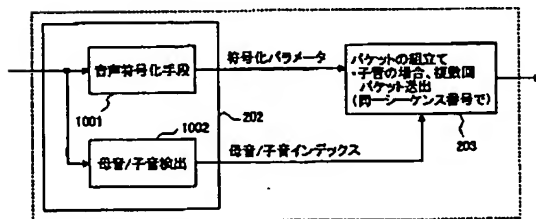
【図 6】

本発明の第1実施例の処理の流れを示す図



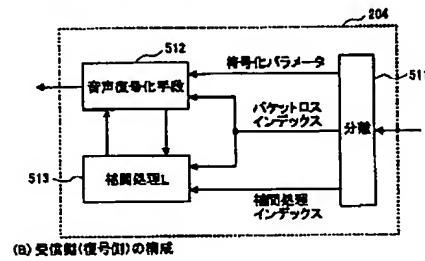
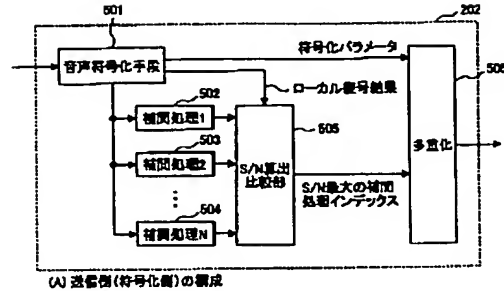
【図 10】

本発明の第3実施例



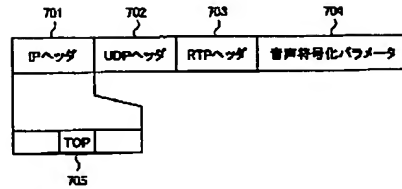
【図 5】

本発明の第1実施例



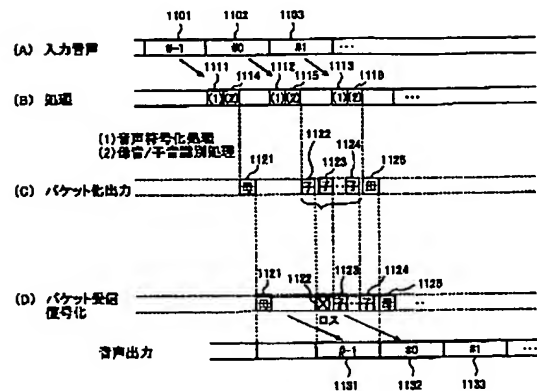
【図 7】

パケットの構成例



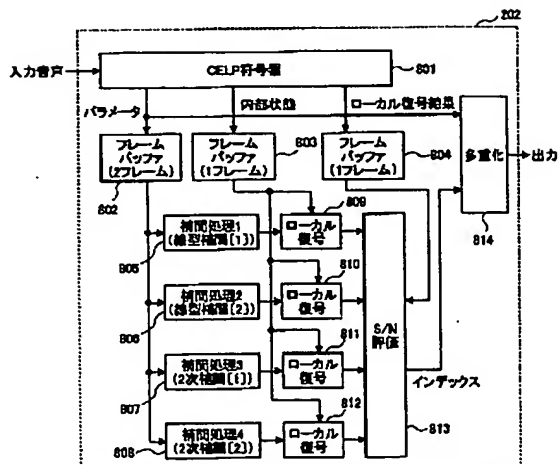
【図 11】

本発明の第3実施例の処理の流れを示す図

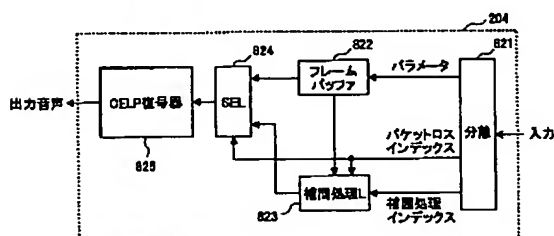


【図8】

本発明の第2実施例



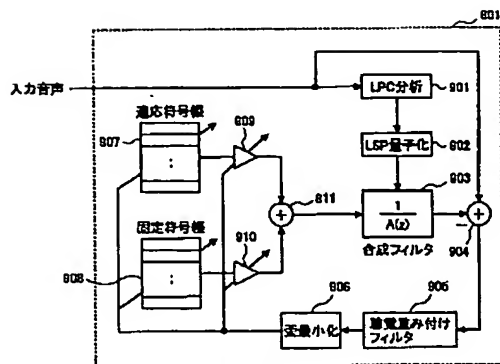
(A) 符号器の構成



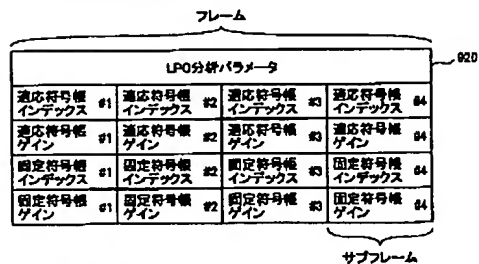
(B) 復号器の構成

【図9】

CELP符号化方式を示す図



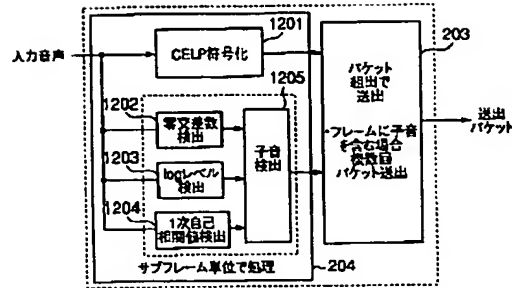
(A) CELP符号器の基本構成



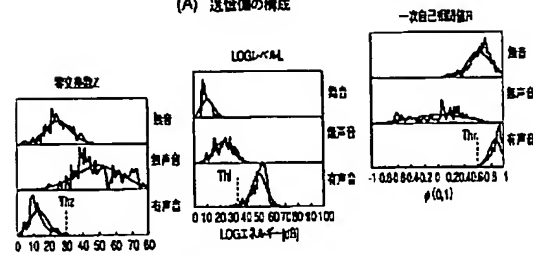
(B) パラメータの伝送タイミング

【図12】

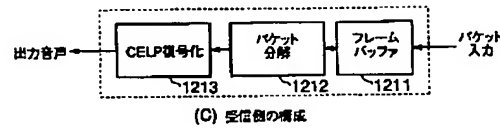
本発明の第4実施例



(A) 送信側の構成



(B) 検出特性分布例



(C) 受信側の構成